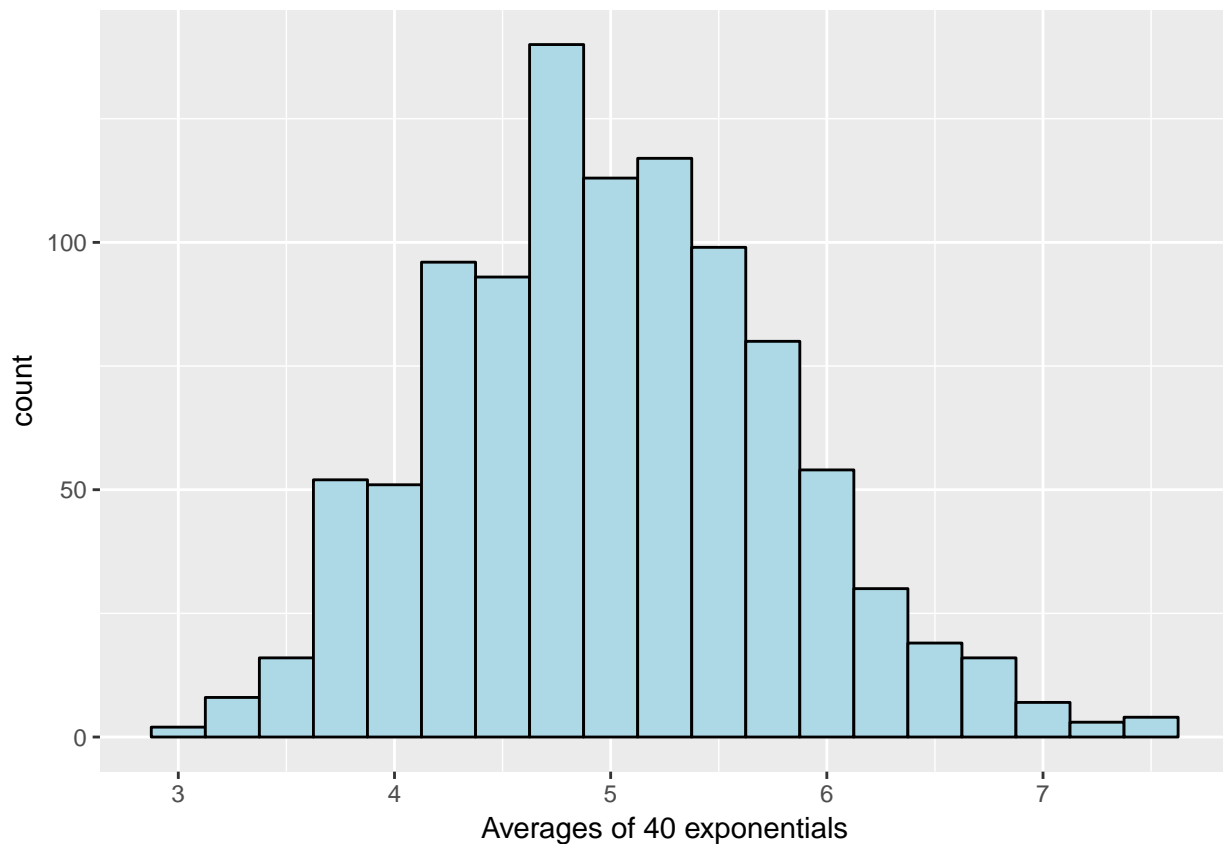# Exponential Distribution Simulation

This project investigates the exponential distribution in R using a simulation. The mean and the standard deviation of an exponential distribution are both 1/lambda, where lambda is the rate parameter.

In this simulation, lambda is set to 0.2, and we are investigating the distribution of averages of 40 exponentials using 1000 simulations. This R code creates a NULL mns object to hold the simulations. It then loops to repeatedly draw 40 exponentials with a lambda of 0.2, take the mean of those 40 exponentials, then store it in mns. The final code outputs a histogram of the means.

```
mns = NULL
for (i in 1:1000) mns = c(mns, mean(rexp(40, 0.2)))
ggplot(data = data.frame(mns), aes(mns)) +
    geom_histogram(col = "black", fill = "light blue", binwidth = 0.25) +
    xlab("Averages of 40 exponentials")
```


1. Sample Mean vs. Theoretical Mean

The sample mean and the theoretical mean are very close. The sample mean of the simulation of 40 exponentials is:

```
mean(mns)
```

```
## [1] 5.006623
```

While the theoretical mean of the distribution is 1/lambda, or 1/0.2:

```r
1/0.2
```

```
## [1] 5
```

2. Sample Variance vs. Theoretical Variance

The sample variance and theoretical variance are also very close in this simulation. The sample variance of the distribution of means is:

```r
sd(mns)^2
```

```
## [1] 0.6160375
```

The theoretical variance of the distribution of means is (1/lambda)^2/n, or:

```r
(1/0.2)^2/40
```

```
## [1] 0.625
```

3. Distribution

Looking at the histogram above gives a relatively good indication that this might be a normal curve, but overlaying a normal curve with the same mean and variance makes it even more clear.

This R code establishes a normal curve adjusted for the mean and standard deviation of the distribution, then plots it along with the histogram from above.

```r
x <- seq(-4,4, length = 1000)*sd(mns) + mean(mns)
y <- dnorm(x, mean(mns), sd(mns))*1000*0.25
xy <- cbind(x, y)

ggplot(data = data.frame(mns), aes(mns)) +
    geom_histogram(col = "black", fill = "light blue", binwidth = 0.25) +
    geom_line(data = data.frame(xy), aes(x=x, y=y)) +
    xlab("Averages of 40 exponentials")
```